

SEQUENCE PROTOCOL

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Werner Lubitz  
(B) ROAD: Schoenborngasse 12/7  
(C) CITY: Wien  
(E) COUNTRY: Austria  
(F) ZIP CODE: 1080

(A) NAME: Uwe Sleytr  
(B) ROAD: Parhamerplatz 10  
(C) CITY: Wien  
(E) COUNTRY: Austria  
(F) ZIP CODE: 1170

(ii) TITLE OF INVENTION: Recombinant expression of  
S-layer proteins

(iii) NUMBER OF SEQUENCES: 9

(iv) COMPUTER READABLE FORM:

(A) DATA CARRIER: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: Patent In Release #1.0, version  
#1.30 (EPA)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3687 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: both  
(D) TOPOLOGY: linear

(vi) INITIAL ORIGIN:

(A) ORGANISM: *Bacillus stearothermophilus*  
(B) STRAIN: PV72

(vii) IMMEDIATE ORIGIN:

(B) CLONE(S): sbsA

(ix) CHARACTERISTIC:

(A) NAME/KEY: CDS  
(B) POSITION: 1..3684

(ix) CHARACTERISTIC:

(A) NAME/KEY: sig peptide  
(B) POSITION: 1..90

SEQUENCE PROTOCOL

Sub B'

(ix) CHARACTERISTIC:

(A) NAME/KEY: mat\_peptide

(B) POSITION: 91..3684

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

10

15

20

25

30

35

40

45

50

55

60

65

ATG	GAT	AGG	AAA	AAA	GCT	GTG	AAA	CTA	GCA	ACA	GCA	AGT	GCT	ATT	GCA	48
Met	Asp	Arg	Lys	Lys	Ala	Val	Lys	Leu	Ala	Thr	Ala	Ser	Ala	Ile	Ala	
-30					-25				-20						-15	
GCA	AGT	GCA	TTT	GTC	GCT	GCA	AAT	CCA	AAC	GCT	TCT	GAA	GCG	GCT	ACA	96
Ala	Ser	Ala	Phe	Val	Ala	Ala	Asn	Pro	Asn	Ala	Ser	Glu	Ala	Ala	Thr	
			-10				-5							1		
GAT	GTA	GCA	ACA	GTA	GTA	AGC	CAA	GCA	AAA	GCA	CAG	TTC	AAA	AAA	GCA	144
Asp	Val	Ala	Thr	Val	Val	Ser	Gln	Ala	Lys	Ala	Gln	Phe	Lys	Lys	Ala	
		5					10					15				
TAC	TAT	ACT	TAC	AGC	CAT	ACA	GTA	ACG	GAA	ACT	GGT	GAA	TTC	CCA	AAC	192
Tyr	Tyr	Thr	Tyr	Ser	His	Thr	Val	Thr	Glu	Thr	Gly	Glu	Phe	Pro	Asn	
		20				25					30					
ATT	AAC	GAT	GTA	TAT	GCT	GAA	TAC	AAC	AAA	GCG	AAA	AAA	CGA	TAC	CGT	240
Ile	Asn	Asp	Val	Tyr	Ala	Glu	Tyr	Asn	Lys	Ala	Lys	Lys	Arg	Tyr	Arg	
35					40					45					50	
GAT	GCG	GTA	GCA	TTA	GTG	AAT	AAA	GCA	GGT	GGC	GCG	AAA	AAA	GAC	GCT	288
Asp	Ala	Val	Ala	Leu	Val	Asn	Lys	Ala	Gly	Gly	Ala	Lys	Lys	Asp	Ala	
				55					60					65		
TAC	TTA	GCT	GAT	TTA	CAA	AAA	GAA	TAT	GAA	ACT	TAC	GTT	TTC	AAA	GCA	336
Tyr	Leu	Ala	Asp	Leu	Gln	Lys	Glu	Tyr	Glu	Thr	Tyr	Val	Phe	Lys	Ala	
			70					75					80			
AAC	CCT	AAA	TCT	GGC	GAA	GCT	CGT	GTA	GCA	ACT	TAC	ATC	GAT	GCT	TAC	384
Asn	Pro	Lys	Ser	Gly	Glu	Ala	Arg	Val	Ala	Thr	Tyr	Ile	Asp	Ala	Tyr	
		85					90					95				
AAC	TAT	GCA	ACA	AAA	TTA	GAC	GAA	ATG	CGC	CAA	GAG	CTA	GAG	GCT	GCT	432
Asn	Tyr	Ala	Thr	Lys	Leu	Asp	Glu	Met	Arg	Gln	Leu	Leu	Glu	Ala	Ala	
		100				105					110					
GTT	CAA	GCA	AAA	GAT	TTA	GAA	AAA	GCA	GAA	CAA	TAC	TAT	CAC	AAA	ATT	480
Val	Gln	Ala	Lys	Asp	Leu	Glu	Lys	Ala	Glu	Gln	Tyr	Tyr	His	Lys	Ile	
115					120					125					130	
CCT	TAT	GAA	ATT	AAA	ACT	CGC	ACA	GTC	ATT	TTA	GAT	CGC	GTA	TAT	GGT	528
Pro	Tyr	Glu	Ile	Lys	Thr	Arg	Thr	Val	Ile	Leu	Asp	Arg	Val	Tyr	Gly	
				135					140					145		
AAA	ACA	ACT	CGT	GAT	TTA	CTT	CGC	TCT	ACA	TTT	AAA	GCA	AAA	GCA	CAA	576
Lys	Thr	Thr	Arg	Asp	Leu	Leu	Arg	Ser	Thr	Phe	Lys	Ala	Lys	Ala	Gln	
			150					155					160			
GAA	CTT	CGC	GAC	AGC	TTA	ATT	TAT	GAT	ATT	ACC	GTT	GCA	ATG	AAA	GCG	624
Glu	Leu	Arg	Asp	Ser	Leu	Ile	Tyr	Asp	Ile	Thr	Val	Ala	Met	Lys	Ala	
		165					170					175				

	CGC	GAA	GTA	CAA	GAC	GCT	GTG	AAA	GCA	GGC	AAT	TTA	GAC	AAA	GCT	AAA	672
	Arg	Glu	Val	Gln	Asp	Ala	Val	Lys	Ala	Gly	Asn	Leu	Asp	Lys	Ala	Lys	
	180						185					190					
5	GCT	GCT	GTT	GAT	CAA	ATC	AAT	CAA	TAC	TTA	CCA	AAA	GTA	ACA	GAT	GCT	720
	Ala	Ala	Val	Asp	Gln	Ile	Asn	Gln	Tyr	Leu	Pro	Lys	Val	Thr	Asp	Ala	
	195					200					205					210	
10	TTC	AAA	ACT	GAA	CTA	ACA	GAA	GTA	GCG	AAA	AAA	GCA	TTA	GAT	GCA	GAT	768
	Phe	Lys	Thr	Glu	Leu	Thr	Glu	Val	Ala	Lys	Lys	Ala	Leu	Asp	Ala	Asp	
					215					220					225		
15	GAA	GCT	GCG	CTT	ACT	CCA	AAA	GTT	GAA	AGT	GTA	AGT	GCG	ATT	AAC	ACT	816
	Glu	Ala	Ala	Leu	Thr	Pro	Lys	Val	Glu	Ser	Val	Ser	Ala	Ile	Asn	Thr	
				230					235					240			
20	CAA	AAC	AAA	GCT	GTT	GAA	TTA	ACA	GCA	GTA	CCA	GTG	AAC	GGA	ACA	CTA	864
	Gln	Asn	Lys	Ala	Val	Glu	Leu	Thr	Ala	Val	Pro	Val	Asn	Gly	Thr	Leu	
			245					250					255				
25	AAA	TTA	CAA	CTT	TCA	GCT	GCT	GCA	AAT	GAA	GAT	ACA	GTA	AAC	GTA	AAT	912
	Lys	Leu	Gln	Leu	Ser	Ala	Ala	Ala	Asn	Glu	Asp	Thr	Val	Asn	Val	Asn	
		260					265					270					
30	ACT	GTA	CGT	ATC	TAT	AAA	GTG	GAC	GGT	AAC	ATT	CCA	TTT	GCC	CTT	AAT	960
	Thr	Val	Arg	Ile	Tyr	Lys	Val	Asp	Gly	Asn	Ile	Pro	Phe	Ala	Leu	Asn	
						280					285					290	
35	ACG	GCA	GAT	GTT	TCT	TTA	TCT	ACA	GAC	GGA	AAA	ACT	ATC	ACT	GTG	GAT	1008
	Thr	Ala	Asp	Val	Ser	Leu	Ser	Thr	Asp	Gly	Lys	Thr	Ile	Thr	Val	Asp	
					295					300					305		
40	GCT	TCA	ACT	CCA	TTC	GAA	AAT	AAT	ACG	GAG	TAT	AAA	GTA	GTA	GTT	AAA	1056
	Ala	Ser	Thr	Pro	Phe	Glu	Asn	Asn	Thr	Glu	Tyr	Lys	Val	Val	Val	Lys	
				310					315					320			
45	GGT	ATT	AAA	GAC	AAA	AAT	GGC	AAA	GAA	TTT	AAA	GAA	GAT	GCA	TTC	ACT	1104
	Gly	Ile	Lys	Asp	Lys	Asn	Gly	Lys	Glu	Phe	Lys	Glu	Asp	Ala	Phe	Thr	
			325				330						335				
50	TTC	AAG	CTT	CGA	AAT	GAT	GCT	GTA	GTT	ACT	CAA	GTG	TTT	GGA	ACT	AAT	1152
	Phe	Lys	Leu	Arg	Asn	Asp	Ala	Val	Val	Thr	Gln	Val	Phe	Gly	Thr	Asn	
		340					345					350					
55	GTA	ACA	AAC	AAC	ACT	TCT	GTA	AAC	TTA	GCA	GCA	GGT	ACT	TTC	GAC	ACT	1200
	Val	Thr	Asn	Asn	Thr	Ser	Val	Asn	Leu	Ala	Ala	Gly	Thr	Phe	Asp	Thr	
						360					365					370	
60	GAC	GAT	ACT	TTA	ACA	GTA	GTA	TTT	GAT	AAG	TTG	TTA	GCA	CCT	GAA	ACT	1248
	Asp	Asp	Thr	Leu	Thr	Val	Val	Phe	Asp	Lys	Leu	Leu	Ala	Pro	Glu	Thr	
					375					380					385		
65	GTA	AAC	AGC	TCG	AAC	GTT	ACT	ATT	ACA	GAT	GTT	GAA	ACT	GGA	AAA	CGC	1296
	Val	Asn	Ser	Ser	Asn	Val	Thr	Ile	Thr	Asp	Val	Glu	Thr	Gly	Lys	Arg	
				390					395					400			
70	ATT	CCA	GTA	ATT	GCA	TCT	ACT	TCT	GGT	TCT	ACA	ATT	ACT	ATT	ACG	TTA	1344
	Ile	Pro	Val	Ile	Ala	Ser	Thr	Ser	Gly	Ser	Thr	Ile	Thr	Ile	Thr	Leu	
			405					410					415				
75	AAA	GAA	GCG	TTA	GTA	ACT	GGT	AAA	CAA	TAT	AAA	CTT	GCT	ATC	AAT	AAT	1392
	Lys	Glu	Ala	Leu	Val	Thr	Gly	Lys	Gln	Tyr	Lys	Leu	Ala	Ile	Asn	Asn	
		420					425					430					
80	GTT	AAA	ACA	TTA	ACT	GGT	TAC	AAT	GCA	GAA	GCT	TAC	GAG	TTA	GTG	TTC	1440
	Val	Lys	Thr	Leu	Thr	Gly	Tyr	Asn	Ala	Glu	Ala	Tyr	Glu	Leu	Val	Phe	
						440					445					450	

CGCTTTT

	ACT	GCA	AAC	GCA	TCA	GCA	CCA	ACT	GTT	GCT	ACC	GCT	CCT	ACT	ACT	TTA	1488
	Thr	Ala	Asn	Ala	Ser	Ala	Pro	Thr	Val	Ala	Thr	Ala	Pro	Thr	Thr	Leu	
				455						460					465		
5	GGT	GGT	ACA	ACT	TTA	TCT	ACT	GGT	TCT	CTT	ACA	ACA	AAT	GTT	TGG	GGT	1536
	Gly	Gly	Thr	Thr	Leu	Ser	Thr	Gly	Ser	Leu	Thr	Thr	Asn	Val	Trp	Gly	
				470					475					480			
10	AAA	TTG	GCT	GGT	GGT	GTG	AAT	GAA	GCT	GGA	ACT	TAT	TAT	CCT	GGT	CTT	1584
	Lys	Leu	Ala	Gly	Gly	Val	Asn	Glu	Ala	Gly	Thr	Tyr	Tyr	Pro	Gly	Leu	
			485					490					495				
15	CAA	TTC	ACA	ACA	ACG	TTT	GCT	ACT	AAG	TTA	GAC	GAA	TCT	ACT	TTA	GCT	1632
	Gln	Phe	Thr	Thr	Thr	Phe	Ala	Thr	Lys	Leu	Asp	Glu	Ser	Thr	Leu	Ala	
		500					505					510					
	GAT	AAC	TTT	GTA	TTA	GTT	GAA	AAA	GAA	TCT	GGT	ACA	GTT	GTT	GCT	TCT	1680
	Asp	Asn	Phe	Val	Leu	Val	Glu	Lys	Glu	Ser	Gly	Thr	Val	Val	Ala	Ser	
	515					520					525					530	
20	GAA	CTA	AAA	TAT	AAT	GCA	GAC	GCT	AAA	ATG	GTA	ACT	TTA	GTG	CCA	AAA	1728
	Glu	Leu	Lys	Tyr	Asn	Ala	Asp	Ala	Lys	Met	Val	Thr	Leu	Val	Pro	Lys	
					535					540					545		
25	GCG	GAC	CTT	AAA	GAA	AAT	ACA	ATC	TAT	CAA	ATC	AAA	ATT	AAA	AAA	GGC	1776
	Ala	Asp	Leu	Lys	Glu	Asn	Thr	Ile	Tyr	Gln	Ile	Lys	Ile	Lys	Lys	Gly	
				550					555					560			
30	TTG	AAG	TCC	GAT	AAA	GGT	ATT	GAA	TTA	GGC	ACT	GTT	AAC	GAG	AAA	ACA	1824
	Leu	Lys	Ser	Asp	Lys	Gly	Ile	Glu	Leu	Gly	Thr	Val	Asn	Glu	Lys	Thr	
			565					570					575				
35	TAT	GAG	TTC	AAA	ACT	CAA	GAC	TTA	ACT	GCT	CCT	ACA	GTT	ATT	AGC	GTA	1872
	Tyr	Glu	Phe	Lys	Thr	Gln	Asp	Leu	Thr	Ala	Pro	Thr	Val	Ile	Ser	Val	
		580					585					590					
40	ACG	TCT	AAA	AAT	GGC	GAC	GCT	GGA	TTA	AAA	GTA	ACT	GAA	GCT	CAA	GAA	1920
	Thr	Ser	Lys	Asn	Gly	Asp	Ala	Gly	Leu	Lys	Val	Thr	Glu	Ala	Gln	Glu	
	595				600						605					610	
	TTT	ACT	GTG	AAG	TTC	TCA	GAG	AAT	TTA	AAT	ACA	TTT	AAT	GCT	ACA	ACC	1968
	Phe	Thr	Val	Lys	Phe	Ser	Glu	Asn	Leu	Asn	Thr	Phe	Asn	Ala	Thr	Thr	
					615					620					625		
45	GTT	TCG	GGT	AGC	ACA	ATC	ACA	TAC	GGT	CAA	GTT	GCT	GTA	GTA	AAA	GCG	2016
	Val	Ser	Gly	Ser	Thr	Ile	Thr	Tyr	Gly	Gln	Val	Ala	Val	Val	Lys	Ala	
				630				635						640			
50	GGT	GCA	AAC	TTA	TCT	GCT	CTT	ACA	GCA	AGT	GAC	ATC	ATT	CCA	GCT	AGT	2064
	Gly	Ala	Asn	Leu	Ser	Ala	Leu	Thr	Ala	Ser	Asp	Ile	Ile	Pro	Ala	Ser	
			645					650					655				
55	GTT	GAA	GCG	GTT	ACT	GGT	CAA	GAT	GGA	ACA	TAC	AAA	GTG	AAA	GTT	GCT	2112
	Val	Glu	Ala	Val	Thr	Gly	Gln	Asp	Gly	Thr	Tyr	Lys	Val	Lys	Val	Ala	
		660				665						670					
60	GCT	AAC	CAA	TTA	GAA	CGT	AAC	CAA	GGG	TAC	AAA	TTA	GTA	GTG	TTC	GGT	2160
	Ala	Asn	Gln	Leu	Glu	Arg	Asn	Gln	Gly	Tyr	Lys	Leu	Val	Val	Phe	Gly	
	675				680						685					690	
	AAA	GGT	GCA	ACA	GCT	CCT	GTT	AAA	GAT	GCT	GCA	AAT	GCA	AAT	ACT	TTA	2208
	Lys	Gly	Ala	Thr	Ala	Pro	Val	Lys	Asp	Ala	Ala	Asn	Ala	Asn	Thr	Leu	
					695					700					705		
65	GCA	ACT	AAC	TAT	ATC	TAT	ACA	TTT	ACA	ACT	GAA	GGT	CAA	GAC	GTA	ACA	2256
	Ala	Thr	Asn	Tyr	Ile	Tyr	Thr	Phe	Thr	Thr	Glu	Gly	Gln	Asp	Val	Thr	
				710					715					720			

	GCA	CCA	ACG	GTT	ACA	AAA	GTA	TTC	AAA	GGT	GAT	TCT	TTA	AAA	GAC	GCT	2304
	Ala	Pro	Thr	Val	Thr	Lys	Val	Phe	Lys	Gly	Asp	Ser	Leu	Lys	Asp	Ala	
			725					730					735				
5	GAT	GCA	GTT	ACT	ACA	CTT	ACG	AAC	GTT	GAT	GCA	GGT	CAA	AAA	TTC	ACT	2352
	Asp	Ala	Val	Thr	Thr	Leu	Thr	Asn	Val	Asp	Ala	Gly	Gln	Lys	Phe	Thr	
		740					745					750					
10	ATC	CAA	TTT	AGC	GAA	GAA	TTA	AAA	ACT	TCT	AGT	GGT	TCT	TTA	GTG	GGT	2400
	Ile	Gln	Phe	Ser	Glu	Glu	Leu	Lys	Thr	Ser	Ser	Gly	Ser	Leu	Val	Gly	
		755				760					765					770	
15	GGC	AAA	GTA	ACT	GTC	GAG	AAA	TTA	ACA	AAC	AAC	GGA	TGG	GTA	GAT	GCT	2448
	Gly	Lys	Val	Thr	Val	Glu	Lys	Leu	Thr	Asn	Asn	Gly	Trp	Val	Asp	Ala	
					775					780					785		
20	GGT	ACT	GGA	ACA	ACT	GTA	TCA	GTT	GCT	CCT	AAG	ACA	GAT	GCA	AAT	GGT	2496
	Gly	Thr	Gly	Thr	Thr	Val	Ser	Val	Ala	Pro	Lys	Thr	Asp	Ala	Asn	Gly	
					790				795					800			
25	AAA	GTA	ACA	GCT	GCT	GTG	GTT	ACA	TTA	ACT	GGT	CTT	GAC	AAT	AAC	GAC	2544
	Lys	Val	Thr	Ala	Ala	Val	Val	Thr	Leu	Thr	Gly	Leu	Asp	Asn	Asn	Asp	
			805					810					815				
30	AAA	GAT	GCG	AAA	TTG	CGT	CTG	GTA	GTA	GAT	AAG	TCT	TCT	ACT	GAT	GGA	2592
	Lys	Asp	Ala	Lys	Leu	Arg	Leu	Val	Val	Asp	Lys	Ser	Ser	Thr	Asp	Gly	
			820				825					830					
35	ATT	GCT	GAT	GTA	GCT	GGT	AAT	GTA	ATT	AAG	GAA	AAA	GAT	ATT	TTA	ATT	2640
	Ile	Ala	Asp	Val	Ala	Gly	Asn	Val	Ile	Lys	Glu	Lys	Asp	Ile	Leu	Ile	
			835			840					845					850	
40	CGT	TAC	AAC	AGC	TGG	AGA	CAC	ACT	GTA	GCT	TCT	GTG	AAA	GCT	GCT	GCT	2688
	Arg	Tyr	Asn	Ser	Trp	Arg	His	Thr	Val	Ala	Ser	Val	Lys	Ala	Ala	Ala	
					855					860					865		
45	GAC	AAA	GAT	GGT	CAA	AAC	GCT	TCT	GCT	GCA	TTC	CCA	ACA	AGC	ACT	GCA	2736
	Asp	Lys	Asp	Gly	Gln	Asn	Ala	Ser	Ala	Ala	Phe	Pro	Thr	Ser	Thr	Ala	
				870					875					880			
50	ATT	GAT	ACA	ACT	AAG	AGC	TTA	TTA	GTT	GAA	TTC	AAT	GAA	ACT	GAT	TTA	2784
	Ile	Asp	Thr	Thr	Lys	Ser	Leu	Leu	Val	Glu	Phe	Asn	Glu	Thr	Asp	Leu	
			885					890					895				
55	GCG	GAA	GTT	AAA	CCT	GAG	AAC	ATC	GTT	GTT	AAA	GAT	GCA	GCA	GGT	AAT	2832
	Ala	Glu	Val	Lys	Pro	Glu	Asn	Ile	Val	Val	Lys	Asp	Ala	Ala	Gly	Asn	
		900					905					910					
60	GCG	GTA	GCT	GGT	ACT	GTA	ACA	GCA	TTA	GAC	GGT	TCT	ACA	AAT	AAA	TTT	2880
	Ala	Val	Ala	Gly	Thr	Val	Thr	Ala	Leu	Asp	Gly	Ser	Thr	Asn	Lys	Phe	
		915				920					925					930	
65	GTA	TTC	ACT	CCA	TCT	CAA	GAA	TTA	AAA	GCT	GGT	ACA	GTT	TAC	TCT	GTA	2928
	Val	Phe	Thr	Pro	Ser	Gln	Glu	Leu	Lys	Ala	Gly	Thr	Val	Tyr	Ser	Val	
					935					940					945		
70	ACA	ATT	GAC	GGT	GTG	AGA	GAT	AAA	GTA	GGT	AAC	ACA	ATC	TCT	AAA	TAC	2976
	Thr	Ile	Asp	Gly	Val	Arg	Asp	Lys	Val	Gly	Asn	Thr	Ile	Ser	Lys	Tyr	
				950					955					960			
75	ATT	ACT	TCG	TTC	AAG	ACT	GTA	TCT	GCG	AAT	CCA	ACG	TTA	TCT	TCA	ATC	3024
	Ile	Thr	Ser	Phe	Lys	Thr	Val	Ser	Ala	Asn	Pro	Thr	Leu	Ser	Ser	Ile	
			965					970					975				
80	AGC	ATT	GCT	GAC	GGT	GCA	GTT	AAC	GTT	GAC	CGT	TCT	AAA	ACA	ATT	ACA	3072
	Ser	Ile	Ala	Asp	Gly	Ala	Val	Asn	Val	Asp	Arg	Ser	Lys	Thr	Ile	Thr	
		980					985					990					

3044-4000

	ATT GAA TTC AGC GAT TCA GTT CCA AAC CCA ACA ATC ACT CTT AAG AAG	3120
	Ile Glu Phe Ser Asp Ser Val Pro Asn Pro Thr Ile Thr Leu Lys Lys	
	995 1000 1005 1010	
5	GCT GAC GGA ACT TCA TTT ACT AAT TAC ACT TTA GTA AAT GTA AAT AAT	3168
	Ala Asp Gly Thr Ser Phe Thr Asn Tyr Thr Leu Val Asn Val Asn Asn	
	1015 1020 1025	
10	GAA AAT AAA ACA TAC AAA ATT GTA TTC CAC AAA GGT GTA ACA CTT GAC	3216
	Glu Asn Lys Thr Tyr Lys Ile Val Phe His Lys Gly Val Thr Leu Asp	
	1030 1035 1040	
15	GAG TTT ACT CAA TAT GAG TTA GCA GTT TCA AAA GAT TTT CAA ACT GGT	3264
	Glu Phe Thr Gln Tyr Glu Leu Ala Val Ser Lys Asp Phe Gln Thr Gly	
	1045 1050 1055	
	ACT GAT ATT GAT AGC AAA GTT ACA TTC ATC ACA GGT TCT GTT GCT ACT	3312
	Thr Asp Ile Asp Ser Lys Val Thr Phe Ile Thr Gly Ser Val Ala Thr	
	1060 1065 1070	
20	GAC GAA GTA AAA CCT GCT CTA GTA GGC GTT GGT TCA TGG AAT GGA ACA	3360
	Asp Glu Val Lys Pro Ala Leu Val Gly Val Gly Ser Trp Asn Gly Thr	
	1075 1080 1085 1090	
25	AGC TAT ACT CAG GAT GCT GCA GCA ACA CGA CTT CGG TCT GTA GCT GAC	3408
	Ser Tyr Thr Gln Asp Ala Ala Ala Thr Arg Leu Arg Ser Val Ala Asp	
	1095 1100 1105	
30	TTC GTT GCG GAG CCA GTT GCC CTT CAA TTC TCA GAA GGT ATC GAT TTA	3456
	Phe Val Ala Glu Pro Val Ala Leu Gln Phe Ser Glu Gly Ile Asp Leu	
	1110 1115 1120	
	ACG AAT GCA ACT GTG ACA GTA ACA AAT ATT ACT GAT GAT AAA ACT GTT	3504
	Thr Asn Ala Thr Val Thr Val Thr Asn Ile Thr Asp Asp Lys Thr Val	
	1125 1130 1135	
35	GAA GTT ATT TCA AAA GAG AGT GTA GAC GCA GAC CAT GAT GCA GGT GCT	3552
	Glu Val Ile Ser Lys Glu Ser Val Asp Ala Asp His Asp Ala Gly Ala	
	1140 1145 1150	
40	ACT AAG GAG ACA TTA GTA ATT AAC ACA GTT ACT CCT TTA GTA CTT GAT	3600
	Thr Lys Glu Thr Leu Val Ile Asn Thr Val Thr Pro Leu Val Leu Asp	
	1155 1160 1165 1170	
45	AAC AGC AAG ACT TAT AAG ATT GTT GTA AGT GGA GTT AAA GAT GCA GCA	3648
	Asn Ser Lys Thr Tyr Lys Ile Val Val Ser Gly Val Lys Asp Ala Ala	
	1175 1180 1185	
50	GGT AAT GTT GCA GAT ACT ATT ACA TTC TAT ATT AAG TAA	3687
	Gly Asn Val Ala Asp Thr Ile Thr Phe Tyr Ile Lys	
	1190 1195	

55 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1228 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Asp Arg Lys Lys Ala Val Lys Leu Ala Thr Ala Ser Ala Ile Ala  
-30 -25 -20 -15

5 Ala Ser Ala Phe Val Ala Ala Asn Pro Asn Ala Ser Glu Ala Ala Thr  
-10 -5 1

Asp Val Ala Thr Val Val Ser Gln Ala Lys Ala Gln Phe Lys Lys Ala  
5 10 15

10 Tyr Tyr Thr Tyr Ser His Thr Val Thr Glu Thr Gly Glu Phe Pro Asn  
20 25 30

Ile Asn Asp Val Tyr Ala Glu Tyr Asn Lys Ala Lys Lys Arg Tyr Arg  
15 35 40 45 50

Asp Ala Val Ala Leu Val Asn Lys Ala Gly Gly Ala Lys Lys Asp Ala  
55 60 65

20 Tyr Leu Ala Asp Leu Gln Lys Glu Tyr Glu Thr Tyr Val Phe Lys Ala  
70 75 80

Asn Pro Lys Ser Gly Glu Ala Arg Val Ala Thr Tyr Ile Asp Ala Tyr  
85 90 95

25 Asn Tyr Ala Thr Lys Leu Asp Glu Met Arg Gln Glu Leu Glu Ala Ala  
100 105 110

Val Gln Ala Lys Asp Leu Glu Lys Ala Glu Gln Tyr Tyr His Lys Ile  
30 115 120 125 130

Pro Tyr Glu Ile Lys Thr Arg Thr Val Ile Leu Asp Arg Val Tyr Gly  
135 140 145

35 Lys Thr Thr Arg Asp Leu Leu Arg Ser Thr Phe Lys Ala Lys Ala Gln  
150 155 160

Glu Leu Arg Asp Ser Leu Ile Tyr Asp Ile Thr Val Ala Met Lys Ala  
40 165 170 175

Arg Glu Val Gln Asp Ala Val Lys Ala Gly Asn Leu Asp Lys Ala Lys  
180 185 190

45 Ala Ala Val Asp Gln Ile Asn Gln Tyr Leu Pro Lys Val Thr Asp Ala  
195 200 205 210

Phe Lys Thr Glu Leu Thr Glu Val Ala Lys Lys Ala Leu Asp Ala Asp  
215 220 225

50 Glu Ala Ala Leu Thr Pro Lys Val Glu Ser Val Ser Ala Ile Asn Thr  
230 235 240

Gln Asn Lys Ala Val Glu Leu Thr Ala Val Pro Val Asn Gly Thr Leu  
245 250 255

55 Lys Leu Gln Leu Ser Ala Ala Ala Asn Glu Asp Thr Val Asn Val Asn  
260 265 270

Thr Val Arg Ile Tyr Lys Val Asp Gly Asn Ile Pro Phe Ala Leu Asn  
60 275 280 285 290

Thr Ala Asp Val Ser Leu Ser Thr Asp Gly Lys Thr Ile Thr Val Asp  
295 300 305

65 Ala Ser Thr Pro Phe Glu Asn Asn Thr Glu Tyr Lys Val Val Val Lys  
310 315 320

SECRET 44460

Gly Ile Lys Asp Lys Asn Gly Lys Glu Phe Lys Glu Asp Ala Phe Thr  
 325 330 335  
 Phe Lys Leu Arg Asn Asp Ala Val Val Thr Gln Val Phe Gly Thr Asn  
 5 340 345 350  
 Val Thr Asn Asn Thr Ser Val Asn Leu Ala Ala Gly Thr Phe Asp Thr  
 355 360 365 370  
 10 Asp Asp Thr Leu Thr Val Val Phe Asp Lys Leu Leu Ala Pro Glu Thr  
 375 380 385  
 Val Asn Ser Ser Asn Val Thr Ile Thr Asp Val Glu Thr Gly Lys Arg  
 390 395 400  
 15 Ile Pro Val Ile Ala Ser Thr Ser Gly Ser Thr Ile Thr Ile Thr Leu  
 405 410 415  
 Lys Glu Ala Leu Val Thr Gly Lys Gln Tyr Lys Leu Ala Ile Asn Asn  
 20 420 425 430  
 Val Lys Thr Leu Thr Gly Tyr Asn Ala Glu Ala Tyr Glu Leu Val Phe  
 435 440 445 450  
 25 Thr Ala Asn Ala Ser Ala Pro Thr Val Ala Thr Ala Pro Thr Thr Leu  
 455 460 465  
 Gly Gly Thr Thr Leu Ser Thr Gly Ser Leu Thr Thr Asn Val Trp Gly  
 470 475 480  
 30 Lys Leu Ala Gly Gly Val Asn Glu Ala Gly Thr Tyr Tyr Pro Gly Leu  
 485 490 495  
 Gln Phe Thr Thr Thr Phe Ala Thr Lys Leu Asp Glu Ser Thr Leu Ala  
 35 500 505 510  
 Asp Asn Phe Val Leu Val Glu Lys Glu Ser Gly Thr Val Val Ala Ser  
 515 520 525 530  
 40 Glu Leu Lys Tyr Asn Ala Asp Ala Lys Met Val Thr Leu Val Pro Lys  
 535 540 545  
 Ala Asp Leu Lys Glu Asn Thr Ile Tyr Gln Ile Lys Ile Lys Lys Gly  
 550 555 560  
 45 Leu Lys Ser Asp Lys Gly Ile Glu Leu Gly Thr Val Asn Glu Lys Thr  
 565 570 575  
 Tyr Glu Phe Lys Thr Gln Asp Leu Thr Ala Pro Thr Val Ile Ser Val  
 50 580 585 590  
 Thr Ser Lys Asn Gly Asp Ala Gly Leu Lys Val Thr Glu Ala Gln Glu  
 595 600 605 610  
 55 Phe Thr Val Lys Phe Ser Glu Asn Leu Asn Thr Phe Asn Ala Thr Thr  
 615 620 625  
 Val Ser Gly Ser Thr Ile Thr Tyr Gly Gln Val Ala Val Val Lys Ala  
 630 635 640  
 60 Gly Ala Asn Leu Ser Ala Leu Thr Ala Ser Asp Ile Ile Pro Ala Ser  
 645 650 655  
 Val Glu Ala Val Thr Gly Gln Asp Gly Thr Tyr Lys Val Lys Val Ala  
 65 660 665 670

30021 4444 1008



	Ala	Asn	Gln	Leu	Glu	Arg	Asn	Gln	Gly	Tyr	Lys	Leu	Val	Val	Phe	Gly
	675					680					685					690
5	Lys	Gly	Ala	Thr	Ala	Pro	Val	Lys	Asp	Ala	Ala	Asn	Ala	Asn	Thr	Leu
					695					700						705
	Ala	Thr	Asn	Tyr	Ile	Tyr	Thr	Phe	Thr	Thr	Glu	Gly	Gln	Asp	Val	Thr
				710					715					720		
10	Ala	Pro	Thr	Val	Thr	Lys	Val	Phe	Lys	Gly	Asp	Ser	Leu	Lys	Asp	Ala
			725					730					735			
	Asp	Ala	Val	Thr	Thr	Leu	Thr	Asn	Val	Asp	Ala	Gly	Gln	Lys	Phe	Thr
	740						745					750				
15	Ile	Gln	Phe	Ser	Glu	Glu	Leu	Lys	Thr	Ser	Ser	Gly	Ser	Leu	Val	Gly
	755					760					765					770
	Gly	Lys	Val	Thr	Val	Glu	Lys	Leu	Thr	Asn	Asn	Gly	Trp	Val	Asp	Ala
20					775					780						785
	Gly	Thr	Gly	Thr	Thr	Val	Ser	Val	Ala	Pro	Lys	Thr	Asp	Ala	Asn	Gly
					790				795					800		
25	Lys	Val	Thr	Ala	Ala	Val	Val	Thr	Leu	Thr	Gly	Leu	Asp	Asn	Asn	Asp
			805					810					815			
	Lys	Asp	Ala	Lys	Leu	Arg	Leu	Val	Val	Asp	Lys	Ser	Ser	Thr	Asp	Gly
	820						825					830				
30	Ile	Ala	Asp	Val	Ala	Gly	Asn	Val	Ile	Lys	Glu	Lys	Asp	Ile	Leu	Ile
	835					840					845					850
	Arg	Tyr	Asn	Ser	Trp	Arg	His	Thr	Val	Ala	Ser	Val	Lys	Ala	Ala	Ala
35					855					860						865
	Asp	Lys	Asp	Gly	Gln	Asn	Ala	Ser	Ala	Ala	Phe	Pro	Thr	Ser	Thr	Ala
				870					875					880		
40	Ile	Asp	Thr	Thr	Lys	Ser	Leu	Leu	Val	Glu	Phe	Asn	Glu	Thr	Asp	Leu
			885					890					895			
	Ala	Glu	Val	Lys	Pro	Glu	Asn	Ile	Val	Val	Lys	Asp	Ala	Ala	Gly	Asn
	900						905					910				
45	Ala	Val	Ala	Gly	Thr	Val	Thr	Ala	Leu	Asp	Gly	Ser	Thr	Asn	Lys	Phe
	915					920					925					930
	Val	Phe	Thr	Pro	Ser	Gln	Glu	Leu	Lys	Ala	Gly	Thr	Val	Tyr	Ser	Val
50					935					940					945	
	Thr	Ile	Asp	Gly	Val	Arg	Asp	Lys	Val	Gly	Asn	Thr	Ile	Ser	Lys	Tyr
				950					955					960		
55	Ile	Thr	Ser	Phe	Lys	Thr	Val	Ser	Ala	Asn	Pro	Thr	Leu	Ser	Ser	Ile
			965					970					975			
	Ser	Ile	Ala	Asp	Gly	Ala	Val	Asn	Val	Asp	Arg	Ser	Lys	Thr	Ile	Thr
	980						985					990				
60	Ile	Glu	Phe	Ser	Asp	Ser	Val	Pro	Asn	Pro	Thr	Ile	Thr	Leu	Lys	Lys
	995					1000					1005					1010
	Ala	Asp	Gly	Thr	Ser	Phe	Thr	Asn	Tyr	Thr	Leu	Val	Asn	Val	Asn	Asn
65					1015					1020						1025

SECRET - 44-2760

(2) INFORMATION FOR SEQ ID NO: 3:

(A) LENGTH: 33 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single strand  
(D) TOPOLOGY: linear

TTAATCGATT CTAGATGGAT AGGAAAAAAG CTG

(2) INFORMATION FOR SEQ ID NO: 4:

(A) LENGTH: 37 base pairs  
(B) TYPE: nucleotide  
(C) STRANDEDNESS: single strand  
(D) TOPOLOGY: linear

ATACCCGGGG GTACGGATCC GATACAGATT TGAGCAA

37

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2766 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(i) INITIAL ORIGIN:

- (A) ORGANISM: Bacillus stearothermophilus
- (B) STRAIN: PV72

(vii) IMMEDIATE ORIGIN:

- (B) CLONE(S): sbsB

(ix) CHARACTERISTIC:

- (A) NAME/KEY: CDS
- (B) POSTION: 1..2763

(ix) CHARACTERISTIC:

- (A) NAME/KEY: sig\_peptide
- (B) POSTION: 1..93

(ix) CHARACTERISTIC:

- (A) NAME/KEY: mat\_peptide
- (B) POSTION: 94..2763

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATG GCT TAT CAA CCT AAG TCT TTT CGC AAG TTT GTT GCG ACA ACT GCA	48
Met Ala Tyr Gln Pro Lys Ser Phe Arg Lys Phe Val Ala Thr Thr Ala	
-31 -30 -25 -20	
ACA GCT GCC ATT GTA GCA TCT GCG GTA GCT CCT GTA GTA TCT GCA GCA	96
Thr Ala Ala Ile Val Ala Ser Ala Val Ala Pro Val Val Ser Ala Ala	
-15 -10 -5 1	
AGC TTC ACA GAT GTT GCG CCG CAA TAT AAA GAT GCG ATC GAT TTC TTA	144
Ser Phe Thr Asp Val Ala Pro Gln Tyr Lys Asp Ala Ile Asp Phe Leu	
5 10 15	
GTA TCA ACT GGT GCA ACA AAA GGT AAA ACA GAA ACA AAA TTC GGC GTT	192
Val Ser Thr Gly Ala Thr Lys Gly Lys Thr Glu Thr Lys Phe Gly Val	
20 25 30	
TAC GAT GAA ATC ACT CGT CTA GAT GCG GCA GTT ATT CTT GCA AGA GTA	240
Tyr Asp Glu Ile Thr Arg Leu Asp Ala Ala Val Ile Leu Ala Arg Val	
35 40 45	

00444-10000

	TTA	AAA	CTA	GAC	GTT	GAC	AAC	GCA	AAA	GAC	GCA	GGC	TTC	ACA	GAT	GTG	288
	Leu	Lys	Leu	Asp	Val	Asp	Asn	Ala	Lys	Asp	Ala	Gly	Phe	Thr	Asp	Val	
	50					55					60					65	
5	CCA	AAA	GAC	CGT	GCA	AAA	TAC	GTC	AAC	GCG	CTT	GTA	GAA	GCT	GGC	GTA	336
	Pro	Lys	Asp	Arg	Ala	Lys	Tyr	Val	Asn	Ala	Leu	Val	Glu	Ala	Gly	Val	
					70					75					80		
10	TTA	AAC	GGT	AAA	GCA	CCT	GGC	AAA	TTT	GGT	GCA	TAC	GAC	CCA	TTA	ACT	384
	Leu	Asn	Gly	Lys	Ala	Pro	Gly	Lys	Phe	Gly	Ala	Tyr	Asp	Pro	Leu	Thr	
				85					90					95			
15	CGC	GTT	GAA	ATG	GCA	AAA	ATC	ATC	GCG	AAC	CGT	TAC	AAA	TTA	AAA	GCT	432
	Arg	Val	Glu	Met	Ala	Lys	Ile	Ile	Ala	Asn	Arg	Tyr	Lys	Leu	Lys	Ala	
			100					105					110				
20	GAC	GAT	GTA	AAA	CTT	CCA	TTC	ACT	GAT	GTA	AAC	GAT	ACA	TGG	GCA	CCA	480
	Asp	Asp	Val	Lys	Leu	Pro	Phe	Thr	Asp	Val	Asn	Asp	Thr	Trp	Ala	Pro	
			115				120					125					
25	TAC	GTA	AAA	GCG	CTT	TAT	AAA	TAC	GAA	GTA	ACC	AAA	AGG	TTA	AAA	CAC	528
	Tyr	Val	Lys	Ala	Leu	Tyr	Lys	Tyr	Glu	Val	Thr	Lys	Arg	Leu	Lys	His	
	130					135					140					145	
30	CAA	CAA	GCT	TCG	GTG	CAT	ACC	AAA	AAC	ATC	ACT	CTG	CGT	GAC	TTT	GCG	576
	Gln	Gln	Ala	Ser	Val	His	Thr	Lys	Asn	Ile	Thr	Leu	Arg	Asp	Phe	Ala	
					150					155					160		
35	CAA	TTT	GTA	TAT	AGA	GCG	GTG	AAT	ATT	AAT	GCA	GTG	CCA	GAA	ATA	GTT	624
	Gln	Phe	Val	Tyr	Arg	Ala	Val	Asn	Ile	Asn	Ala	Val	Pro	Glu	Ile	Val	
				165					170					175			
40	GAA	GTA	ACT	GCG	GTT	AAT	TCG	ACT	ACA	GTG	AAA	GTA	ACA	TTC	AAT	ACG	672
	Glu	Val	Thr	Ala	Val	Asn	Ser	Thr	Thr	Val	Lys	Val	Thr	Phe	Asn	Thr	
			180					185					190				
45	CAA	ATT	GCT	GAT	GTT	GAT	TTC	ACA	AAT	TTT	GCT	ATC	GAT	AAC	GGT	TTA	720
	Gln	Ile	Ala	Asp	Val	Asp	Phe	Thr	Asn	Phe	Ala	Ile	Asp	Asn	Gly	Leu	
		195					200					205					
50	ACT	GTT	ACT	AAA	GCA	ACT	CTT	TCT	CGT	GAT	AAA	AAA	TCC	GTA	GAG	GTT	768
	Thr	Val	Thr	Lys	Ala	Thr	Leu	Ser	Arg	Asp	Lys	Lys	Ser	Val	Glu	Val	
						215					220					225	
55	GTG	GTA	AAT	AAA	CCG	TTT	ACT	CGT	AAT	CAG	GAA	TAT	ACA	ATT	ACA	GCG	816
	Val	Val	Asn	Lys	Pro	Phe	Thr	Arg	Asn	Gln	Glu	Tyr	Thr	Ile	Thr	Ala	
					230					235					240		
60	ACA	GGC	ATT	AAA	AAT	TTA	AAA	GGC	GAG	ACC	GCT	AAG	GAA	TTA	ACT	GGT	864
	Thr	Gly	Ile	Lys	Asn	Leu	Lys	Gly	Glu	Thr	Ala	Lys	Glu	Leu	Thr	Gly	
				245					250					255			
65	AAG	TTT	GTT	TGG	TCT	GTT	CAA	GAT	GCG	GTA	ACT	GTT	GCA	CTA	AAT	AAT	912
	Lys	Phe	Val	Trp	Ser	Val	Gln	Asp	Ala	Val	Thr	Val	Ala	Leu	Asn	Asn	
			260					265					270				
70	AGT	TCG	CTT	AAA	GTT	GGA	GAG	GAA	TCT	GGT	TTA	ACT	GTA	AAA	GAT	CAG	960
	Ser	Ser	Leu	Lys	Val	Gly	Glu	Glu	Ser	Gly	Leu	Thr	Val	Lys	Asp	Gln	
			275				280					285					
75	GAT	GGC	AAA	GAT	GTT	GTA	GGT	GCT	AAA	GTA	GAA	CTT	ACT	TCT	TCT	AAT	1008
	Asp	Gly	Lys	Asp	Val	Val	Gly	Ala	Lys	Val	Glu	Leu	Thr	Ser	Ser	Asn	
						295					300					305	
80	ACT	AAT	ATT	GTT	GTA	GTT	TCA	AGT	GGC	GAA	GTA	TCA	GTA	TCT	GCT	GCT	1056
	Thr	Asn	Ile	Val	Val	Val	Ser	Ser	Gly	Glu	Val	Ser	Val	Ser	Ala	Ala	
					310					315					320		

00144-1000

65 AGT GAC ACA ACA CCA ACT TCA ACT AAA ACA ATT ACT GTT AAT GTA GTA 1872  
Ser Asp Thr Thr Pro Thr Ser Thr Lys Thr Ile Thr Val Asn Val Val  
580 585 590

	Asn	Val	Lys	Ala	Asp	GAC	GCT	ACA	CCA	GTA	GGG	TTA	GAT	ATT	GTA	GCA	CCT	
	595							600					605					1920
5	TCT Ser 610	AAA Lys	ATT Ile	GAT Asp	GTA Val	AAT Asn	GCT Ala	CCA Pro	AAC Asn	ACT Thr	GCT Ala	TCT Ser	ACT Thr	GCA Ala	GAT Asp	GTT Val		1968
10	GAT Asp	TTT Phe	ATA Ile	AAT Asn	TTC Phe	GAA Glu	AGT Ser	GTT Val	GAG Glu	ATT Ile	TAC Tyr	ACA Thr	CTC Leu	GAT Asp	TCA Ser	AAT Asn		2016
15	GGT Gly	AGA Arg	CGT Arg	CAA Gln	AAA Lys	AAA Lys	GTT Val	ACT Thr	CCA Pro	ACT Thr	GCA Ala	ACT Thr	ACA Thr	CTT Leu	GTA Val	GGT Gly		2064
20	ACA Thr	AAA Lys	AAA Lys	AAA Lys	AAA Lys	AAA Lys	GTT Val	AAT Asn	GGG Gly	AAT Asn	GTA Val	TTA Leu	CAA Gln	TTC Phe	AAG Lys	GGG Gly		2112
25	AAC Asn	GAA Glu	GAA Glu	TTA Leu	ACG Thr	CTA Leu	TCA Ser	ACT Thr	TCT Ser	TCT Ser	AGT Ser	ACA Thr	GGA Gly	AAC Asn	GTA Val	GAT Asp		2160
30	GGA Gly	ACA Thr	GCA Ala	GAA Glu	GGA Gly	ATG Met	ACA Thr	AAA Lys	CGT Arg	ATT Ile	CCA Pro	GGG Gly	AAA Lys	TAT Tyr	ATC Ile	AAC Asn		2208
35	TCT Ser	GCA Ala	AGT Ser	GTA Val	CCT Pro	GCC Ala	AGT Ser	GCA Ala	ACA Thr	GTA Val	GCA Ala	ACA Thr	AGT Ser	CCT Pro	GTT Val	ACT Thr		2256
40	GTA Val	AAG Lys	CTT Leu	AAT Asn	TCA Ser	AGT Ser	GAT Asp	AAT Asn	GAT Asp	TTA Leu	ACA Thr	TTT Phe	GAA Glu	GAA Glu	TTA Leu	ATA Ile		2304
45	TTC Phe	GGT Gly	GTA Val	ATT Ile	GAC Asp	CCT Pro	ACA Thr	CAA Gln	TTA Leu	GTC Val	AAA Lys	GAT Asp	GAA Glu	GAC Asp	ATC Ile	AAC Asn		2352
50	GAA Glu	TTT Phe	ATT Ile	GCA Ala	GTT Val	TCA Ser	AAA Lys	GCG Ala	GCT Ala	AAA Lys	AAT Asn	GAT Asp	GGA Gly	TAT Tyr	TTG Leu	TAT Tyr		2400
55	AAT Asn	AAA Lys	CCG Pro	CTT Leu	GTA Val	ACG Thr	GTT Val	AAA Lys	GAT Asp	GCA Ala	TCA Ser	GGA Gly	AAA Lys	GTT Val	ATT Ile	CCA Pro		2448
60	ACA Thr	GGT Gly	GCA Ala	AAT Asn	GTT Val	TAC Tyr	GGT Gly	CTA Leu	AAT Asn	CAT His	GAT Asp	GCA Ala	ACT Thr	AAC Asn	GGA Gly	AAC Asn		2496
65	ATT Ile	TGG Trp	TTT Phe	GAT Asp	GAG Glu	GAA Glu	CAA Gln	GCT Ala	GGC Gly	TTA Leu	GCT Ala	AAA Lys	AAA Lys	TTT Phe	AGT Ser	GAT Asp		2544
70	GTA Val	CAT His	TTT Phe	GAT Asp	GTT Val	GAT Asp	TTT Phe	TCA Ser	TTA Leu	ACT Thr	AAC Asn	GTT Val	GTA Val	AAA Lys	ACT Thr	GGT Gly		2592
75	AGC Ser	GGT Gly	ACA Thr	GTT Val	TCT Ser	TCA Ser	TCG Ser	CCA Pro	TCA Ser	TTA Leu	TCT Ser	GAC Asp	GCA Ala	ATT Ile	CAA Gln	CTT Leu		2640
80	ACT Thr	AAT Asn	TCA Ser	GGC Gly	GAT Asp	GCA Ala	GTA Val	TCG Ser	TTT Phe	ACA Thr	TTA Leu	GTT Val	ATC Ile	AAA Lys	TCA Ser	ATT Ile		2688

TAT GTT AAA GGC GCA GAT AAA GAT GAT AAT AAC TTA CTT GCA GCC CCT 2736  
Tyr Val Lys Gly Ala Asp Lys Asp Asp Asn Asn Leu Leu Ala Ala Pro  
870 875 880

5 GTT TCT GTC AAT GTG ACT GTG ACA AAA TAA 2766  
Val Ser Val Asn Val Thr Val Thr Lys  
885 890

10 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 921 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

20  
Met Ala Tyr Gln Pro Lys Ser Phe Arg Lys Phe Val Ala Thr Thr Ala  
25 -31 -30 -25 -20  
Thr Ala Ala Ile Val Ala Ser Ala Val Ala Pro Val Val Ser Ala Ala  
-15 -10 -5 1  
30 Ser Phe Thr Asp Val Ala Pro Gln Tyr Lys Asp Ala Ile Asp Phe Leu  
5 10 15  
Val Ser Thr Gly Ala Thr Lys Gly Lys Thr Glu Thr Lys Phe Gly Val  
20 25 30  
35 Tyr Asp Glu Ile Thr Arg Leu Asp Ala Ala Val Ile Leu Ala Arg Val  
35 40 45  
Leu Lys Leu Asp Val Asp Asn Ala Lys Asp Ala Gly Phe Thr Asp Val  
40 50 55 60 65  
Pro Lys Asp Arg Ala Lys Tyr Val Asn Ala Leu Val Glu Ala Gly Val  
70 75 80  
45 Leu Asn Gly Lys Ala Pro Gly Lys Phe Gly Ala Tyr Asp Pro Leu Thr  
85 90 95  
Arg Val Glu Met Ala Lys Ile Ile Ala Asn Arg Tyr Lys Leu Lys Ala  
100 105 110  
50 Asp Asp Val Lys Leu Pro Phe Thr Asp Val Asn Asp Thr Trp Ala Pro  
115 120 125  
Tyr Val Lys Ala Leu Tyr Lys Tyr Glu Val Thr Lys Arg Leu Lys His  
55 130 135 140 145  
Gln Gln Ala Ser Val His Thr Lys Asn Ile Thr Leu Arg Asp Phe Ala  
150 155 160  
60 Gln Phe Val Tyr Arg Ala Val Asn Ile Asn Ala Val Pro Glu Ile Val  
165 170 175  
Glu Val Thr Ala Val Asn Ser Thr Thr Val Lys Val Thr Phe Asn Thr  
180 185 190

00117447-10000

	Gln	Ile	Ala	Asp	Val	Asp	Phe	Thr	Asn	Phe	Ala	Ile	Asp	Asn	Gly	Leu
	195						200					205				
5	Thr	Val	Thr	Lys	Ala	Thr	Leu	Ser	Arg	Asp	Lys	Lys	Ser	Val	Glu	Val
	210					215					220					225
	Val	Val	Asn	Lys	Pro	Phe	Thr	Arg	Asn	Gln	Glu	Tyr	Thr	Ile	Thr	Ala
					230					235					240	
10	Thr	Gly	Ile	Lys	Asn	Leu	Lys	Gly	Glu	Thr	Ala	Lys	Glu	Leu	Thr	Gly
				245					250					255		
	Lys	Phe	Val	Trp	Ser	Val	Gln	Asp	Ala	Val	Thr	Val	Ala	Leu	Asn	Asn
			260					265					270			
15	Ser	Ser	Leu	Lys	Val	Gly	Glu	Glu	Ser	Gly	Leu	Thr	Val	Lys	Asp	Gln
		275					280					285				
20	Asp	Gly	Lys	Asp	Val	Val	Gly	Ala	Lys	Val	Glu	Leu	Thr	Ser	Ser	Asn
		290				295					300					305
	Thr	Asn	Ile	Val	Val	Val	Ser	Ser	Gly	Glu	Val	Ser	Val	Ser	Ala	Ala
					310					315					320	
25	Lys	Val	Thr	Ala	Val	Lys	Pro	Gly	Thr	Ala	Asp	Val	Thr	Ala	Lys	Val
				325					330					335		
	Thr	Leu	Pro	Asp	Gly	Val	Val	Leu	Thr	Asn	Thr	Phe	Lys	Val	Thr	Val
			340					345					350			
30	Thr	Glu	Val	Pro	Val	Gln	Val	Gln	Asn	Gln	Gly	Phe	Thr	Leu	Val	Asp
		355				360						365				
35	Asn	Leu	Ser	Asn	Ala	Pro	Gln	Asn	Thr	Val	Ala	Phe	Asn	Lys	Ala	Glu
		370				375					380					385
	Lys	Val	Thr	Ser	Met	Phe	Ala	Gly	Glu	Thr	Lys	Thr	Val	Ala	Met	Tyr
					390					395					400	
40	Asp	Thr	Lys	Asn	Gly	Asp	Pro	Glu	Thr	Lys	Pro	Val	Asp	Phe	Lys	Asp
				405					410					415		
	Ala	Thr	Val	Arg	Ser	Leu	Asn	Pro	Ile	Ile	Ala	Thr	Ala	Ala	Ile	Asn
			420					425					430			
45	Gly	Ser	Glu	Leu	Leu	Val	Thr	Ala	Asn	Ala	Gly	Gln	Ser	Gly	Lys	Ala
		435					440					445				
50	Ser	Phe	Glu	Val	Thr	Leu	Lys	Asp	Asn	Thr	Lys	Arg	Thr	Phe	Thr	Val
		450				455					460					465
	Asp	Val	Lys	Lys	Asp	Pro	Val	Leu	Gln	Asp	Ile	Lys	Val	Asp	Ala	Thr
					470					475					480	
55	Ser	Val	Lys	Leu	Ser	Asp	Glu	Ala	Val	Gly	Gly	Gly	Glu	Val	Glu	Gly
				485					490					495		
	Val	Asn	Gln	Lys	Thr	Ile	Lys	Val	Ser	Ala	Val	Asp	Gln	Tyr	Gly	Lys
			500					505					510			
60	Glu	Ile	Lys													



Asp Ser Gly Asn Ser Ala Thr Asp Gln Phe Val Val Val Ala Thr Lys  
550 555 560

5 Asp Lys Ile Val Asn Gly Lys Val Glu Val Lys Tyr Phe Lys Asn Ala  
565 570 575

Ser Asp Thr Thr Pro Thr Ser Thr Lys Thr Ile Thr Val Asn Val Val  
580 585 590

10 Asn Val Lys Ala Asp Ala Thr Pro Val Gly Leu Asp Ile Val Ala Pro  
595 600 605

Ser Lys Ile Asp Val Asn Ala Pro Asn Thr Ala Ser Thr Ala Asp Val  
610 615 620 625

15 Asp Phe Ile Asn Phe Glu Ser Val Glu Ile Tyr Thr Leu Asp Ser Asn  
630 635 640

Gly Arg Arg Gln Lys Lys Val Thr Pro Thr Ala Thr Thr Leu Val Gly  
645 650 655

20 Thr Lys Lys Lys Lys Lys Val Asn Gly Asn Val Leu Gln Phe Lys Gly  
660 665 670

25 Asn Glu Glu Leu Thr Leu Ser Thr Ser Ser Ser Thr Gly Asn Val Asp  
675 680 685

Gly Thr Ala Glu Gly Met Thr Lys Arg Ile Pro Gly Lys Tyr Ile Asn  
690 695 700 705

30 Ser Ala Ser Val Pro Ala Ser Ala Thr Val Ala Thr Ser Pro Val Thr  
710 715 720

35 Val Lys Leu Asn Ser Ser Asp Asn Asp Leu Thr Phe Glu Glu Leu Ile  
725 730 735

Phe Gly Val Ile Asp Pro Thr Gln Leu Val Lys Asp Glu Asp Ile Asn  
740 745 750

40 Glu Phe Ile Ala Val Ser Lys Ala Ala Lys Asn Asp Gly Tyr Leu Tyr  
755 760 765

Asn Lys Pro Leu Val Thr Val Lys Asp Ala Ser Gly Lys Val Ile Pro  
770 775 780 785

45 Thr Gly Ala Asn Val Tyr Gly Leu Asn His Asp Ala Thr Asn Gly Asn  
790 795 800

50 Ile Trp Phe Asp Glu Glu Gln Ala Gly Leu Ala Lys Lys Phe Ser Asp  
805 810 815

Val His Phe Asp Val Asp Phe Ser Leu Thr Asn Val Val Lys Thr Gly  
820 825 830

55 Ser Gly Thr Val Ser Ser Ser Pro Ser Leu Ser Asp Ala Ile Gln Leu  
835 840 845

Thr Asn Ser Gly Asp Ala Val Ser Phe Thr Leu Val Ile Lys Ser Ile  
850 855 860 865

60 Tyr Val Lys Gly Ala Asp Lys Asp Asp Asn Asn Leu Leu Ala Ala Pro  
870 875 880

65 Val Ser Val Asn Val Thr Val Thr Lys  
885 890

SECRET

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 498 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CCCATGGACC CGTCCAAGGA CTCCAAAGCT CAGGTTTCTG CAGCCGAAGC TGGTATCACT 60  
GGCACCTGGT ATAACCAACT GGGGTCGACT TTCATTGTGA CCGCTGGTGC GGACGGAGCT 120  
CTGACTGGCA CCTACGAATC TGCGGTTGGT AACGCAGAAT CCCGCTACGT ACTGACTGGC 180  
CGTTATGACT CTGCACCTGC CACCGATGGC TCTGGTACCG CTCTGGGCTG GACTGTGGCT 240  
TGA AAAACA ACTATCGTAA TGCGCACAGC GCCACTACGT GGTCTGGCCA ATACGTTGGC 300  
GGTGCTGAGG CTCGTATCAA CACTCAGTGG CTGTTAACAT CCGGCACTAC CGAAGCGAAT 360  
GCATGGAAAT CGACACTAGT AGGTCATGAC ACCTTTACCA AAGTTAAGCC TTCTGCTGCT 420  
AGCATTGATG CTGCCAAGAA AGCAGGCGTA AACAACGGTA ACCCTCTAGA CGCTGTTCAG 480  
CAATAATAAG GATCCGGG 498

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8

TTCATCGTAA ACGCCGAATT TTGTTTCTG 29

INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: single strand
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AGGGAAATAT ATCAACTCTG CAAGTG 26